Seismic Tomography Using Variational Inference Methods

Xin Zhang and Andrew Curtis
University of Edinburgh, School of Geosciences, Edinburgh, United Kingdom of Great Britain and Northern Ireland (x.zhang2@ed.ac.uk)

In a variety of geoscientific applications we require maps of subsurface properties together with the corresponding maps of uncertainties to assess their reliability. Seismic tomography is a method that is widely used to generate those maps. Since tomography is significantly nonlinear, Monte Carlo sampling methods are often used for this purpose, but they are generally computationally intractable for large data sets and high-dimensionality parameter spaces. To extend uncertainty analysis to larger systems, we introduce variational inference methods to conduct seismic tomography. In contrast to Monte Carlo sampling, variational methods solve the Bayesian inference problem as an optimization problem yet still provide fully nonlinear, probabilistic results. This is achieved by minimizing the Kullback-Leibler (KL) divergence between approximate and target probability distributions within a predefined family of probability distributions.

We introduce two variational inference methods: automatic differential variational inference (ADVI) and Stein variational gradient descent (SVGD). In ADVI a Gaussian probability distribution is assumed and optimized to approximate the posterior probability distribution. In SVGD a smooth transform is iteratively applied to an initial probability distribution to obtain an approximation to the posterior probability distribution. At each iteration the transform is determined by seeking the steepest descent direction that minimizes the KL-divergence.

We apply the two variational inference methods to 2D travel time tomography using both synthetic and real data, and compare the results to those obtained from two different Monte Carlo sampling methods: Metropolis-Hastings Markov chain Monte Carlo (MH-McMC) and reversible jump Markov chain Monte Carlo (rj-McMC). The results show that ADVI provides a biased approximation because of its Gaussian approximation, whereas SVGD produces more accurate approximations to the results of MH-McMC. In comparison rj-McMC produces smoother mean velocity models and lower standard deviations because the parameterization used in rj-McMC (Voronoi cells) imposes prior restrictions on the pixelated form of models: all pixels within each Voronoi cell have identical velocities. This suggests that the results of rj-McMC need to be interpreted in the light of the specific prior information imposed by the parameterization. Both variational methods estimate the posterior distribution at significantly lower computational cost, provided that gradients of parameters with respect to data can be calculated efficiently. We therefore expect that the methods can be applied fruitfully to many other types of geophysical inverse problems.